

12-3-2013

Data Management Planning: NIH Data Sharing Policy and Resources

Lisa Zilinski

Purdue University, ldz@andrew.cmu.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_fspres



Part of the [Information Literacy Commons](#)

Recommended Citation

Zilinski, Lisa, "Data Management Planning: NIH Data Sharing Policy and Resources" (2013). *Libraries Faculty and Staff Presentations*. Paper 41.
http://docs.lib.purdue.edu/lib_fspres/41

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Webinar Series: Data Literacy for Librarians, December 3, 2013

Data Management Planning

NIH Data Sharing Policy and Resources

Lisa D. Zilinski

*Data Services Specialist, Assistant Professor
Purdue University Libraries*

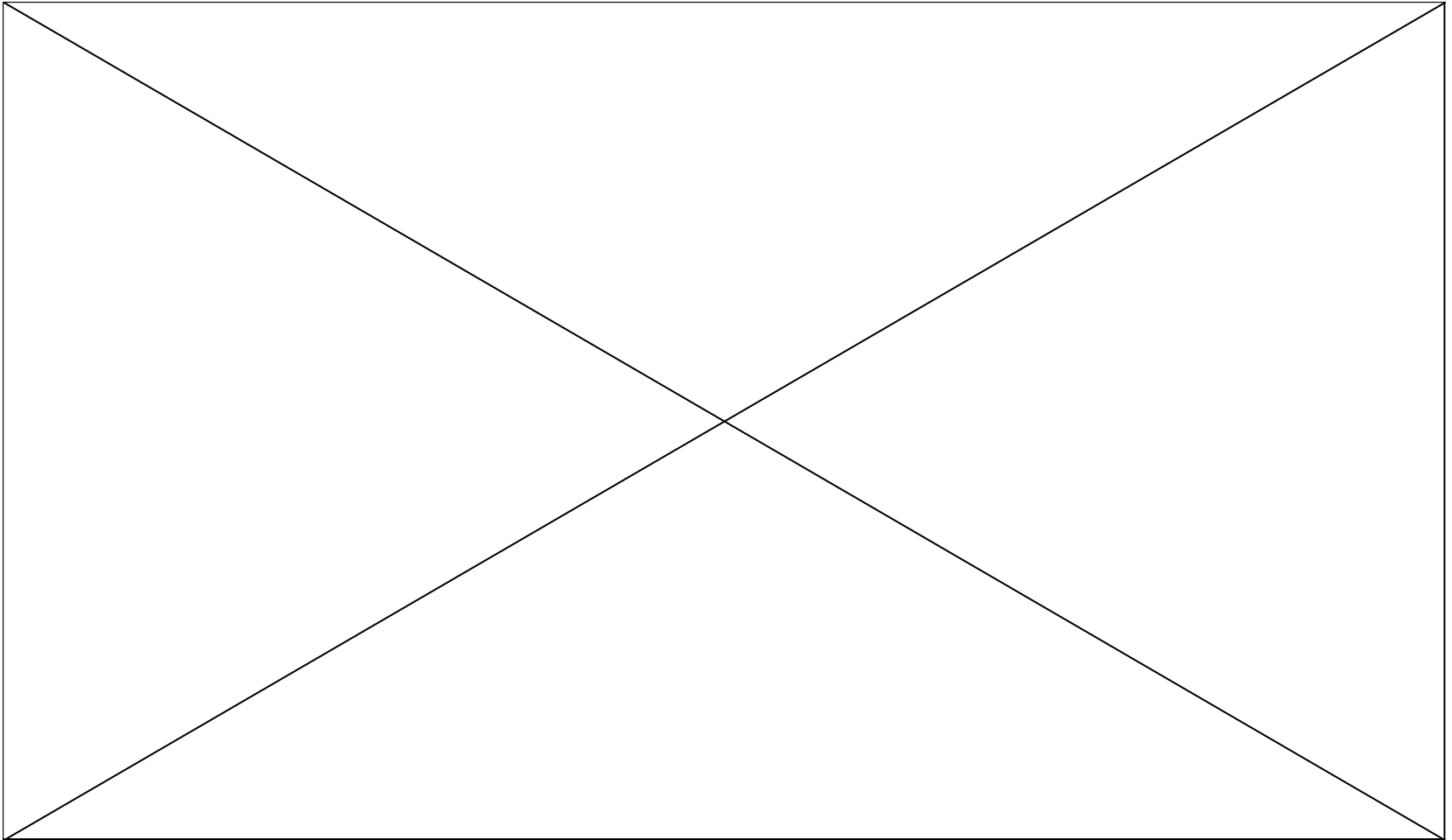


Agenda

- Introduction to Data Sharing
- NIH Policy
- Elements of Data Sharing and Management Plans
- The NSF Data Management Plan
- Resources
- Wrap-Up and Questions



Scenario



Hanson, K., Surkis, A., & Yacobucci, L. (2012, August 3). Data Sharing and Management Snafu in 3 Short Acts. NYU Health Science Libraries. [Video file]. Retrieved from <http://youtu.be/N2zK3sAtr-4>

Data Sharing

- The National Institutes of Health (NIH) believes “that data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human health.” ¹
- Data Sharing Goals ²
 - ▶ *Reinforces open scientific inquiry*
 - ▶ *Encourages diversity of analysis and opinion*
 - ▶ *Promotes new research*
 - ▶ *Allows reuse for testing new/alternative hypotheses and methods*
 - ▶ *Supports studies on data collection methods/measurements*
 - ▶ *Facilitates education of new researchers*
 - ▶ *Enables further exploration of research topics*
 - ▶ *Permits creation of new datasets through the combination of multiple source datasets*

The Library

■ Additional benefits of data sharing

- ▶ *Discovery*
- ▶ *Context*
- ▶ *Security*
- ▶ *Impact*



■ Some benefits of data management planning

- ▶ *Preservation*
- ▶ *Documentation*
- ▶ *Meeting funding mandates*

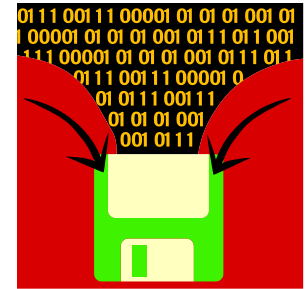
NIH Policy on Data Sharing

- Applications for grants of \$500,000 or more are required to include a data sharing plan.
- This should include a description of how the final research data will be shared.
- If the data will not be shared, an explanation must be included in the application.
- NIH expects timely release and sharing of final research data.
- Data sharing plans are required for all grant applications that include genome-wide association studies, regardless of the proposed budget.

Elements of a Data Sharing Plan

Questions to consider when writing an NIH data sharing plan: ³

- ▶ *What data will be shared?*
- ▶ *Who will have access to the data?*
- ▶ *Where will the data to be shared be located?*
- ▶ *When will the data be shared?*
- ▶ *How will researchers locate and access the data?*



Describe Your Data

- Types of data
- Unique data
- Individual-level or raw data
- Aggregate data
- Quality control methods
- Documentation (metadata, schema)
- Data standards and/or vocabularies
- Data formats



Data Access

- General public
- Restrictions to access
- Criteria for restricting access
- Sensitive data access
- Data sharing agreements
- Who will manage access?



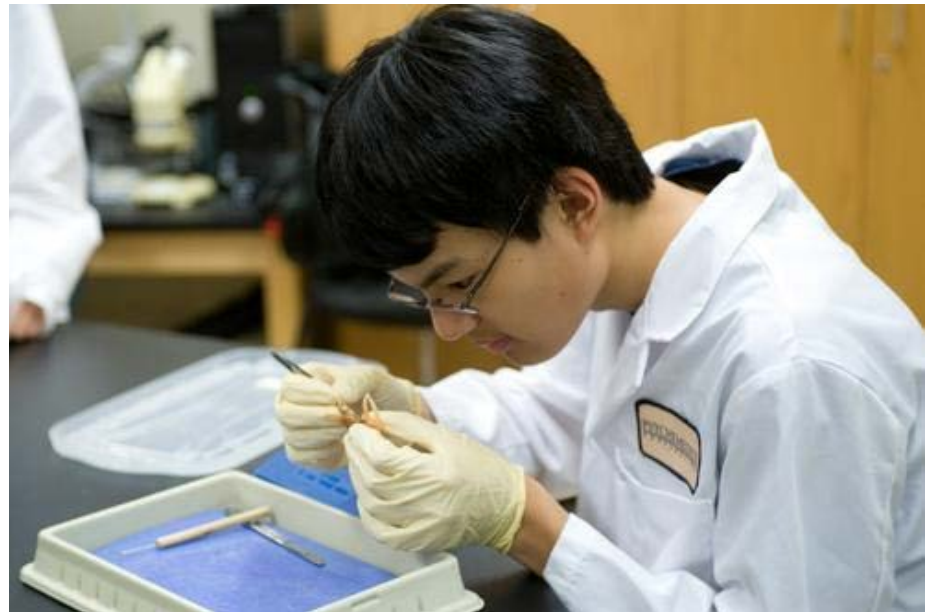
Data Storage

- Existing databases, repositories, enclaves, and archives
- Develop new repository
- Data distribution
- Institutional vs. disciplinary



Data Availability

- Schedule release of data
- Incremental data releases
- Policies for data releases
- Maintenance and access of data



Locate and Access Data

- Registries, repositories, indexes, publications
- Linking and cross-references
- Access methods



Sample DSPs

4

Example Data Sharing Plan for FOA-XX-XXXX

What data that will be shared:

I will share phenotypic data associated with the collected samples by depositing these data at _____ which is an NIH-funded repository. Genotype data will be shared by depositing these data at _____. Additional data documentation and de-identified data will be deposited for sharing along with phenotypic data, which includes demographics, family history of XXXXXX disease, and diagnosis, consistent with applicable laws and regulations. I will comply with the NIH GWAS Policy and the funding IC's existing policies on sharing data on XXXXXX disease genetics to include secondary analysis of data resulting from a genome wide association study through the repository. Meta-analysis data and associated phenotypic data, along with data content, format, and organization, will be available at _____. Submitted data will confirm with relevant data and terminology standards.

Who will have access to the data:

I agree that data will be deposited and made available through _____ which is an NIH-funded repository, and that these data will be shared with investigators working under an institution with a Federal Wide Assurance (FWA) and could be used for secondary study purposes such as finding genes that contribute to process of XXXXXX. I agree that the names and Institutions of persons either given or denied access to the data, and the bases for such decisions, will be summarized in the annual progress report. Meta-analysis data and associated phenotypic data, along with data content, format, and organization, will be made available to investigators through _____.

Where will the data be available:

I agree to deposit and maintain the phenotypic data, and secondary analysis of data (if any) at _____, which is an NIH-funded repository and that the repository has data access policies and procedures consistent with NIH data sharing policies.

When will the data be shared:

I agree to deposit genetic outcome data into _____ repository as soon as possible but no later than within one year of the completion of the funded project period for the parent award or upon acceptance of the data for publication, or public disclosure of a submitted patent application, whichever is earlier.

How will researchers locate and access the data:

I agree that I will identify where the data will be available and how to access the data in any publications and presentations that I author or co-author about these data, as well as acknowledge the repository and funding source in any publications and presentations. As I will be using _____, which is an NIH-funded repository, this repository has policies and procedures in place that will provide data access to qualified researchers, fully consistent with NIH data sharing policies and applicable laws and regulations.

5

Sharing of data generated by this project is an essential part of our proposed activities and will be carried out in several different ways. We would wish to make our results available both to the community of scientists interested in [this disease] and the biology of [its causative agent] to avoid unintentional duplication of research. Conversely, we would welcome collaboration with others who could make use of the vaccine assessment protocols developed in [the project].

Our plan includes the following:

Presentations at national scientific meetings. From the projects, it is expected that approximately four presentations at national meetings would be appropriate. There is an annual [Disease] Study Group meeting, of which the PI is secretary. This one-day meeting of interested persons presents new information on a variety of topics related to [the disease]. It is expected that the investigators from this [project] will be active participants of this focused group.

Annual lectureship. A lectureship has brought to the University distinguished scientists and clinicians whose areas of expertise were relevant to those interested in [the disease]. Lecturers have been [list of names]. Visiting lecturers will be scheduled to interact with the investigators of the project as appropriate with their specific areas of expertise which will provide an opportunity for members to present their work to the visitor.

Newsletter. The [disease interest group] publishes a newsletter which currently has a circulation of [number]. The newsletter's intent is to disseminate new information regarding [the disease]. The activities and discoveries of [the project] will be allocated 20% of the newsletter's coverage.

Web site of the Interest Group. The [interest group] currently maintains a Web site where information [about the disease] is posted. Summaries of the scientific presentation from the [quarterly project] meetings will be posted on this Web site, written primarily for a general audience. [Link to Web site]

Annual [Disease] Awareness week. Beginning this fall during the week of [date], the [interest group] will be sponsoring a [Disease] Awareness week. As part of that program, there will be a research poster display with discussions. In future years, [the project investigators] will be active participants in this program.

SAGE Library Data. [This project] will generate data from several SAGE libraries. It is our explicit intention that these data will be placed in a readily accessible public database. All efforts will be made to rapidly release data through publication of results as quickly as it is possible to analyze the experiments. Data used in publications will be released in a timely manner. SAGE data will be made accessible through a public site that allows querying as has been set up for a similar project. This site can be accessed at [link to Web site].

The NSF Data Management Plan (DMP)

National Science Foundation (NSF) Data Management Plan (DMP) Requirements ⁶

- Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled “Data Management Plan”. This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results , and may include:
 1. *the **types of data**, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;*
 2. *the **standards** to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);*
 3. *policies for **access and sharing** including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;*
 4. *policies and provisions for **re-use**, re-distribution, and the production of derivatives; and*
 5. *plans for **archiving** data, samples, and other research products, and for preservation of access to them.*

Data Sharing and Management Resources

- Disciplinary Metadata Directory



<http://www.dcc.ac.uk/resources/metadata-standards>

- DMPTool



<https://dmp.cdlib.org/>

- Databib



<http://databib.org/>

- DataCite



<http://www.datacite.org/>

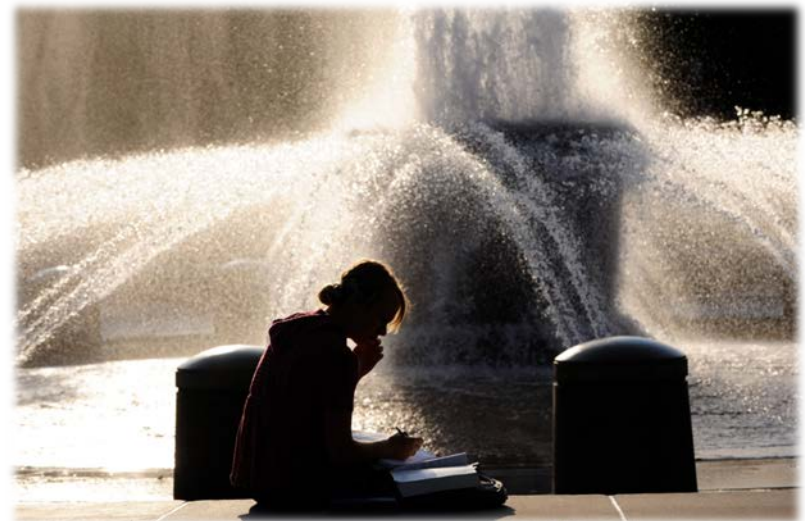
Questions?

Thank you.

Lisa D. Zilinski

lzilinski@purdue.edu

<http://blogs.lib.purdue.edu/lzilinski/>



References and Resources

1. National Institutes of Health. (2003, February 26). *Final NIH Statement on Sharing Research Data*. Retrieved from <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>
2. National Institutes of Health. (2003, March 5). *NIH Data Sharing Policy and Implementation Guidance*. Retrieved from http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm
3. National Institutes of Health. (2009, December 30). *Key Elements to Consider in Preparing a Data Sharing Plan Under NIH Extramural Support*. Retrieved from http://grants.nih.gov/grants/sharing_key_elements_data_sharing_plan.pdf
4. NIH Office of Extramural Research (OER). (2010, August 31). Example Plan addressing Key Elements for a Data Sharing Plan under NIH Extramural Support. Retrieved from http://grants.nih.gov/grants/sharing_example_data_sharing_plan.doc
5. The National Institute of Allergy and Infectious Diseases (NIAID) (2011, October 4). Sample Data Sharing Plan. Retrieved from <http://www.niaid.nih.gov/researchfunding/tool/pages/datasharingex.aspx>
6. National Science Foundation. (2013, January 1). Proposal Preparation Instructions: Plans for data management and sharing. Retrieved from http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg_2.jsp#dmp